No 2015-12 – July

Working Paper

Firm Level Allocative Inefficiency: Evidence from France

Lionel Fontagné & Gianluca Santoni

Highlights

- Firm level misallocation is here defined as the wedge (gap) between the marginal return and marginal cost of a production input.
- A significant increase in the average labour gap for French manufacturing firms is observed from the early 2000s.
- Agglomeration economies positively affect labour allocation at the firm (plant) level, possibly through better matching employer-employees.
- Firms inefficiencies are propagating along Input-Output linkages.

Abstract

A large portion of productivity differentials among locations is related to density. Firms located in denser areas are more productive due to agglomeration economies (Combes *et al.*, 2012). We provide in this paper an explanation of such economies: lower input misallocation. The distribution of resources among heterogeneous firms has relevant consequences on allocative efficiency and denser areas provide a more favorable environment for dynamic matching between employers and employees. Using a methodology proposed by Petrin and Sivadasan (2013) we are able to assess the degree of resource misallocation among firms within sectors for each of the 96 French "Départements". Based on firm-level productivity estimates, we identify in the gap between the value of the marginal product and marginal input price the output loss due to inefficiencies in inputs allocation. Over the period 1993- 2007 the average gap at firm level is around 10 thousands euro, showing a relevant increase starting from the early 2000s. Importantly, firms misallocations are lower in denser areas, suggesting that the matching mechanism is playing a role in explaining the productivity premium of agglomerated locations.

Keywords

Misallocation, Productivity, Firm Level Data.

JEL

D24, L25, O47.

Working Paper

CEPII

CEPII (Centre d'Etudes Prospectives et d'Informations Internationales) is a French institute dedicated to producing independent, policy-oriented economic research helpful to understand the international economic environment and challenges in the areas of trade policy, competitiveness, macroeconomics, international finance and growth.

CEPII Working Paper
Contributing to research in international economics

Editorial Director:
Sébastien Jean

Production:
Laure Boivin

No ISSN: 1293-2574

CEPII
113, rue de Grenelle
75007 Paris
+33 1 53 68 55 00

www.cepii.fr
Press contact: presse@cepii.fr

RESEARCH AND EXPERTISE ON THE WORLD ECONOMY

Firm Level Allocative Inefficiency: Evidence from France[1]

Lionel Fontagné* and Gianluca Santoni†

[1]This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 320197 (Mapcompete project). The views expressed in this publication are the authors' alone, and do not necessarily reflect the views of the European Commission. Lionel Fontagné acknowledges support from the Bank of France. We are grateful to Antoine Berthou, Gilbert Cette, Matthieu Crozet, Ana Margarida Fernandes, Claire Lelarge, Thierry Mayer and Fabienne Rosenwald for helpful suggestions. We are also indebted to Irene Iodice for excellent research assistance.

*Paris School of Economics – Université Paris I and CEPII,(lionel.fontagne@univ-paris1.fr)
†CEPII,(gianluca.santoni@cepii.fr)

Introduction

Denser areas are more productive. This can be firstly due to selection, as only the most productive firms can locate in more competitive environments. This can also be due to agglomeration economies, associated with a better access to a variety of inputs, or the circulation of ideas. Combes et al. (2012) have shown that firms located in denser areas are on average 9.7% more productive with respect to those located in a less dense environment. Their findings suggest that the main driver for such differentials is not selection (i.e. tougher competition inducing less productive firms to exit the market) but agglomeration economies; the latter being generated by three main mechanisms: higher availability of services, infrastructures and public goods (sharing), thicker labor market (matching), technology spillovers (learning)[2].

A relevant feature of firms' productivity distribution is a sizeable and persistent heterogeneity (i.e. dispersion), even when productivity is computed within narrowly defined sectors.[3] Syverson (2004) reports for the US a Total Factor Productivity (TFP) ratio of 1.92 among firms at 90 percentile and 10 percentile of industry distribution: within a narrow defined sector, most productive firms are able to produce almost twice the output of less productive ones, with the same amount of inputs. The degree of misallocation is even higher in China and India, the gain in TFP by achieving the same allocative efficiency as the US would be between 30-50% for China, and as much as 40-60% for India, while the increase in output would be almost two times higher.[4] Such dispersion in firm level outcomes implies that micro-economic behavior does matter for aggregate ones: individual idiosyncratic shocks do not simply average out in the aggregation but they significantly shape the overall outcomes. The evolution of several

[2] Following the classification from Duranton and Puga (2004).

[3] The large variability at firm level is not confined to TFP; for example, sales growth rates in US show a standard deviation of about 50% (Davis et al., 2007), that translates for one third of the firms into an expected growth of more than 60% and for another third to an expected decline of more than 40%. High variability in firms' productivity, sales, entry and exit rates suggests that allocation of resources plays an important role: notwithstanding the more structural employment shifts, the capacity of churning to drive resources towards the most efficient firms is conducive to aggregate performance.

[4] See Hsieh and Klenow (2009). US productivity naturally displays gaps and a degree of misallocation, the distribution is used just as a control group.

macro-economic aggregates, such as productivity, value added, employment and investment are then closely related to what happens at the micro-level. The "granular" hypothesis affirms that large firms idiosyncrasies affect aggregate GDP fluctuations and, through general equilibrium channels, all other firms as well (Gabaix, 2011).[5]

A key driver of productivity dispersion is the resource allocation easiness. Empirical literature, in fact, confirms this critical property: resources (production inputs) do not flow freely from low to high productive firms, even if more efficient firms are the most likely to survive in the markets. Generally, reallocation of economic activity at firm level tends to benefit high productive (low cost) producers, resulting in an aggregate improvement; but several factors may hamper this continuous flow of resources from less to more efficient firms: business cycles[6], labor and capital rigidity, regulation environment and competition. Consequently, from a macroeconomic perspective a large portion of cross-country productivity differentials are imputable to input misallocation: with heterogeneous firms, the distribution of resources among them has significant consequences on both allocation efficiency and aggregate outcome.[7] The usual approach to measuring the degree of efficiency in resource allocation across countries is based on the covariance among firms' size and productivity. If resources were allocated purely randomly such covariance would be zero; conversely the higher the covariance the more efficient resources are allocated across firms (Bartelsman et al., 2009).[8] Market rigidity, distorted regulations and other frictions may weaken the correlation with fundamentals. In this vein, the empirical evidence reported in CompNet (Berthou and Sandoz, 2014)[9] shows that over the period 2003-2007 the

[5]The law of large numbers no longer applies if the distribution of firms' sales departs from normality and displays "fat-tail".

[6]Lazear and Spletzer (2012) show that labor reallocation seems to be more conspicuous during expansionary periods than recessions.

[7]See Hsieh and Klenow (2009), Syverson (2014), Dhingra and Morrow (2014).

[8]The procedure, following Olley and Pakes (1996), uses the covariance between firm size and productivity within sectors to assess the efficiency of input allocation. Note that this is the static version of allocative efficiency, in a cross-section framework; see Haltiwanger (2011) for a discussion on static and dynamic allocative measures.

[9]Competitiveness Research Network, is composed of economists from the 28 national central banks of the European Union (EU) and the European Central Bank; international organizations (World Bank, OECD, EU Commission), universities and think-tanks, as well as non-European Central Banks (Argentina and Peru) and organizations (US International Trade Commission). The objective of CompNet is to develop a more consistent analytical framework for assessing competitiveness, allowing for a better correspondence between determinants and outcomes.

distribution of inputs across European countries could improve significantly. The covariance between labor productivity and firm size reaches 0.2 for Hungary and Spain, meaning that in those countries labor allocation is about 20% more efficient than the random allocation benchmark; a similar analysis for US shows a correlation of about 50%[10].

From a microeconomic perspective, resource misallocation implies that more efficient firms tend to be smaller than their optimal size while less efficient ones tends to be bigger than their optimum production scale. The dispersion of revenue-based productivity (the product of physical productivity and a firm's output price) is revealing the degree of resource misallocation so-defined (Hsieh and Klenow, 2009). The rationale is that, without distortions, revenue-based productivity should be equal for all firms in the same sector. Alternatively, one can look at the difference between the marginal product in value of each factor and its cost for the firm (Petrin and Sivadasan, 2013). Such difference is a *gap*, which measures the degree of resource misallocation among firms, within sectors. It measures the extent to which firms do not fully optimize. Following this line of reasoning, we can shed light on the channels through with agglomerated locations, within an economy, are more productive, instead of looking at cross-country differences in the efficiency of resource allocation. Firms in denser areas – notwithstanding the distortions present in the whole economy, such as labor market rigidities – may indeed match with more productive and better paid workers. But if one tackles the *difference* between wage and the marginal product in value, a better matching should anyway reduce the gap between the two observed at firm level. Using administrative data for the universe of legal units operating in the French manufacturing sector over the period 1993-2007, we show in this paper that this mechanism is present: resource misallocation among firms, within sectors, is lower in denser *Départements*.

We also confirm that sectoral linkages act as a complementary channel through which microeconomic shocks may generate a "cascade effect" (Acemoglu et al., 2012). A localized sectoral

[10]Results from Bartelsman et al. (2013) report relatively higher covariance for European countries, ranging from 15-38 %, confirming a sizeable efficiency gap with respect to US benchmark.

shock propagates to the whole economy through intermediate supplies linkages. Firms and sectoral interconnections are a vehicle propagating the fluctuations to the whole economy.

The rest of the paper is organized as follows. The methodology inspired from Petrin and Sivadasan (2013) used in this paper is detailed in Section 1. The French data is described in Section 2. TFP estimation strategy is described in Section 2.1. The value of Labor Gaps is computed in Section 3, both at the sector level (Section 3.1) and its aggregate evolution (Section 3.2). In Section 4, we assess the dynamics of labor gaps, controlling for firm characteristics. The last section concludes.

1. Measuring Resource Allocation at Firm Level

Given the highlighted interconnections between micro and macroeconomic forces, the distribution of resource may have significant effect on productivity and per capita income. As noted before, input market imperfections (or distortional regulations) can create an incentive for less productive firms to produce beyond their optimal sizes and at the same time hinder most efficient firms. The main consequence is that the economy is producing less than currently available resources would allow, only due to an inefficient distribution of them. In order to evaluate the impact of a change in labor market regulation in Chile, Petrin and Sivadasan (2013) propose a new methodology to assess the degree of resource misallocation at the firm level.

The following empirical work relies on Petrin and Sivasadan (PS) methodology. Their approach, based on plant-level productivity estimates, aims to define the output loss due to inefficiencies in inputs allocation, as well as the impact of policies change at both firm and aggregate level. The key concept of firm specific "mis-allocation" is referring to the *"gap among the value of the marginal product and marginal input price"* (Petrin and Sivadasan, 2013). Such gaps are computed at firm level using the estimated coefficients from usual Total Factor Productivity (TFP) analysis and can be further aggregated at sector or spatial level. Moreover since those gaps are expressed in monetary terms the direct aggregation gives the amount of lost output due to the induced distortion in resource distribution across firms.

The economic intuition behind this approach is that, under perfect competition, an input's value of marginal return should be equated to its marginal cost. A wedge between marginal return and marginal cost is a signal that firms are not fully optimizing, curbing aggregate output. The estimation of the gap from firm-level data starts from a Cobb-Douglas production function for firm i at time t as the following:

$$q_{it} = \beta_l l_{it} + \beta_k k_{it} + \omega_{it} + \varepsilon_{it} \tag{1}$$

Where q_{it} denotes value added, l_{it} the number of employees and k_{it} the fixed capital stock. All the series are in logs and expressed in real terms.[11]

The error term is made of two components: ω_{it} that represents a Hicks-neutral productivity shock (observed by the firm but not by the econometrician) and ε_{it} that is uncorrelated with the input choice (unobservable to the firm and to the econometrician). The main complication here is that ω_{it} will affect input decision at the firm level, inducing a simultaneity bias for the production function estimation. The economic rationale relies on the fact that present investments will be productive only in the next period and a representative firm will choose how much to invest only after observing its current productivity level (see Section 2.1 for a detailed discussion).

Estimating Equation (1) provides a measure of firm's i production efficiency – the difference between the observed and predicted level of output. Essentially, a firm is more productive with respect to another in the same sector if it can produce more output using the same level of inputs[12]. In order to set a benchmark output level we need to estimate the marginal return (i.e. marginal product) of each input for the representative firm within each industry. In what follows we focus on labor marginal productivity (but it can be generalized to any input). The marginal

[11] Our empirical exercise is using industry price deflators from the French National Statistical Institute (INSEE).
[12] Or reaching the same level of output using less input.

product of labor is given as the marginal increment in output per unit change in labor:

$$\begin{aligned}\frac{\partial Q_{it}}{\partial L} &= \beta_l e^{\upsilon_{it}} L_{it}^{\beta_l - 1} K_{it}^{\beta_k} \\ &= \beta_l \frac{Q_{it}}{L_{it}}\end{aligned} \tag{2}$$

Where $\upsilon_{it} = \omega_{it} + \varepsilon_{it}$. Once the marginal product is recovered from the production function estimation – Equation (2) – the value of the marginal product of labor is given by just multiplying the marginal product by the firm level output price.

$$VMP_{it}^{l} = P_{kt}\left(\beta_l \frac{Q_{it}}{L_{it}}\right) \tag{3}$$

Since output prices at firm level are generally not available (or only for a sub-sample of surveyed firms), we use industry specific price index, P_{kt}. Clearly, using industry prices is a cause for concern, see for example Foster et al. (2008), due to the risk of introducing a measurement error in prices. However, as noted by Petrin and Sivadasan (2013), marginal products of inputs, i.e. β_z, are still consistent if the deviation of the plant level price from the industry price is not systematically correlated with the input levels. Moreover, they suggest to condition the calculation of VMP on ω_{it} to control for measurement error. In this case our measure of the marginal product of labor is given by: $\beta_l \frac{Q_{it} e^{\omega_{it}}}{L_{it} e^{\upsilon_{it}}}$ [13], giving more weight to firms with a lower unexpected productivity shock, we use this robustness in all the empirical applications. Finally the degree of resource misallocation at firm level, the revenue to cost gap, is given by:

$$G_{it}^{l} = \mid VMP_{it}^{l} - w_{it} \mid \tag{4}$$

Where w_{it} represents the wage of the marginal worker for firm i[14]. To ease comparability over

[13]Where $\upsilon_{it} = \omega_{it} + \varepsilon_{it}$. In order to isolate the unexpected error term we follow PS approach. We derive expected output $\widehat{q_{it}}$ level by estimating a regression of value added on variable inputs and a polynomial function of capital and material inputs, where $\widehat{q_{it}}$ represents the value of output net of ε_{it}.

[14]Since we do not observe the salary paid to the marginal employee, we use the average wage as proxy. Wage

time, the value of G^l_{it} has been deflated using the Consumer Price Index[15], the value of G^l_{it} in absolute terms expresses the increase in value added induced by an optimal reallocation of labor. In a setting where resources are allocated optimally and there are no frictions in the input markets, all firms will demand labor up-until the expected marginal return will equate the marginal cost closing the gap. In reality there are several reasons why an economy could depart from such equilibrium: hiring and firing costs, capital adjustment costs, taxes but also mark-ups and management practices. According to Petrin and Sivadasan (2013) the social optimum is reached when all gaps are equal to zero, while an efficient allocation of labor implies that gaps are equated across firms (Syverson, 2011).

2. Data and TFP estimation

The evaluation of input allocation is performed using balance sheet firm level data to retrieve TFP estimations, from which we derive the marginal contribution of production inputs. Then, using firm (or industry) specific input prices, it is possible to derive a monetary value of the allocation inefficiencies at firm level.

The main source of firm level data is the French BRN[16] dataset obtained from the fiscal administration. It contains balance-sheet information collected from the firms' tax forms; along with detailed information on the firms' balance sheets, including total, domestic, and export sales, value added, as well as many cost items including the wage bill, materials expenditures, and so on, as well as the sectors and the region in which the firm operates.

The dataset covers the period 1993-2007 giving a very detailed representation of the aggregate economy. The fact that the information come from tax authorities, then, ensures an overall very high quality of the data.

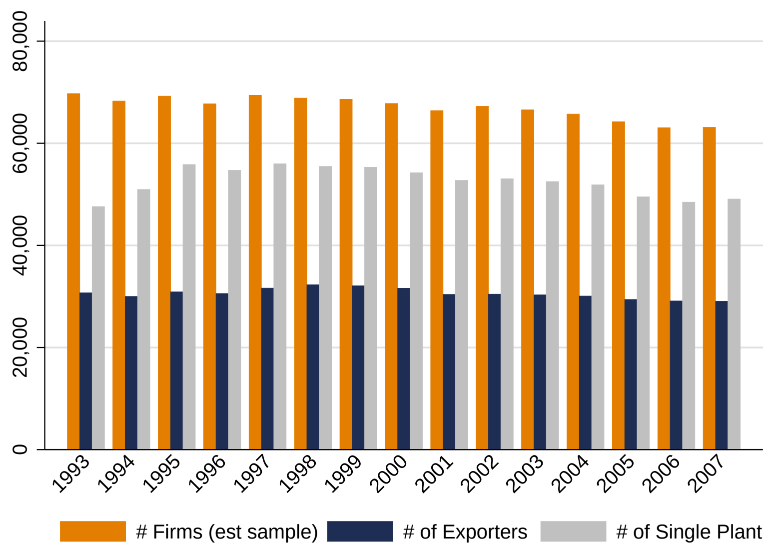
After excluding implausible observations, namely those reporting negative or zero values for our

includes salary and tax allowances.

[15]Results are robust to the use of GDP deflator instead of CPI.

[16]BRN stands for Bénéfice Réel Normal, the normal tax regime for French firms.

Figure 1 – Distribution of Firms in the estimation sample, number of exporters and single-plant firms (manufacturing only)



variables of interest and cleaning the data from potential outliers[17], we end up with an unbalanced panel of more than 138 thousand firms for the manufacturing sector[18]. As reported in Figure 1, while most of the firms in the sample have only one production plant, still the fraction of multi-plant firms in the sample is around 20 percent and is growing over time. Interestingly half of multi-plant firms have their establishments located within the same department (see Table ST5 in the Appendix), suggesting the optimal geographical aggregation for the input misallocation indicator. The share of exporters on the other hand is remarkably stable over time, suggesting that our empirical evidence should not be driven by sample compositional effects. Indeed, it is well known from the empirical literature in international economics that exporters are significantly different on many dimensions with respect to non-exporters (see Bernard et al. (2007) and Wagner (2012) for a recent survey).

[17]We exclude observations with a growth rate of TFP variables – value added, fixed capital, material inputs and services, above/below the 99th/1st percentile of the relative distribution. We also make sure that firm balance sheets cover 12 months. Results are robust to change in the thresholds. As robustness check we also exclude firms with a sales values less than 100 or 750 thousand euro respectively and our main findings remain unaffected.

[18]We limit the analysis to the manufacturing sector only to ease the interpretation of TFP estimation coefficients as marginal products; the underline methodology however can be applied to other industries as well.

2.1. TFP estimation

In order to assess input gaps the first step is to compute firm-level TFP. Our measure of TFP is computed using the Wooldridge (2009) implementation of the Levinsohn and Petrin (2003) algorithm using material inputs as proxy for technology shocks[19] and considering labor as freely adjustable (variable) input and capital as fixed.

The semi-parametric estimator by Levinsohn and Petrin (2003) – LP – extends the methodology of Olley and Pakes (1996), who originally suggested the use of investments as proxy to avoid the problem of simultaneity between technology shock ω_{it} and input choice in a two stage estimation procedure[20]. LP instead suggest to use raw materials as proxy variable for ω_{it} mainly because investments are a valid proxy only if they adjust smoothly to productivity shocks (Petrin et al., 2004), but also because intermediate goods tend to be reported with a higher frequency in firms' balance sheets.

In the two-stage procedure, the demand for materials is represented as a function of state variables: productivity (un-observed) and capital $m_{it} = g(k_{it}, \omega_{it})$. Under the hypothesis that $g(.)$ is invertible, productivity itself can be expressed as $\omega_{it} = f(k_{it}, m_{it})$. In this way productivity is approximated by a function of observable variables such as capital (k_{it}) and intermediates (m_{it}). The two-step estimation proceeds as follows: in the first stage are identified the coefficients for variable inputs while the coefficients for state variables are recovered in the second stage, where $f(.)$ is usually approximated by an n-order polynomial in k_{it} and m_{it}.[21]

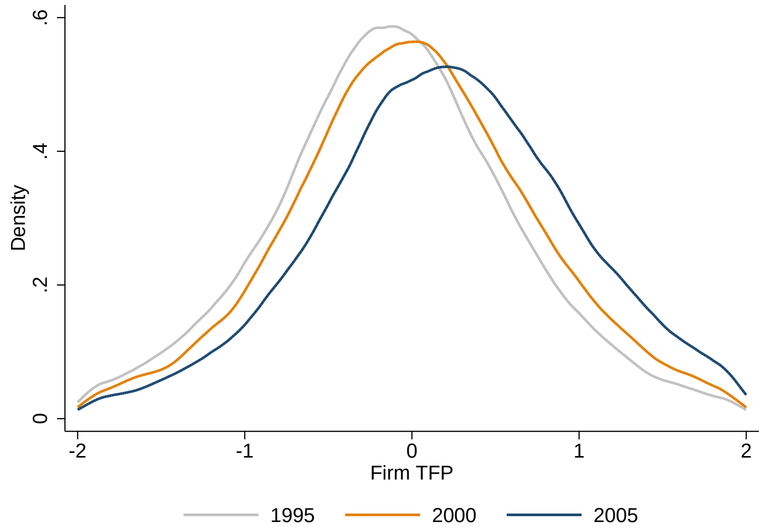
[19]Wooldridge suggests implementing the Levinsohn and Petrin approach in a GMM framework that ensures more efficiency since it takes into account the potential contemporaneous error correlation of the two stages as well as heteroskedasticity and/or serial correlation.

[20]This is the main reason why estimators that ignore such correlation produce inconsistent results, as OLS for example.

[21]See Van Biesebroeck (2007) for a detailed discussion on the different methodologies to estimate productivity (underlying assumptions and drawbacks). Using simulated data he also provides a sensitivity analysis of five estimators to factor price heterogeneity, measurement error and technology differences. Simulation results show that with measurement error or heterogeneous production technology, GMM estimator provides very robust productivity estimates (both for levels and growth rates). Moreover, even in the absence of such distortions, it provides reliable results also if (at least part of) productivity differences are persistent over time.

It is worth noting that our findings are robust to different estimation methodologies, namely: the semi-parametric two stage LP estimator as well as the GMM implementation with labor as a fixed input. The latter case is particularly relevant since it assumes explicitly the existence of labor market frictions[22].

Figure 2 – Manufacturing Firms Productivity Distribution, ω_{it}



Note: The Graph reports the distribution of manufacturing firm total factor productivity for selected years. Pooled distribution has been standardized (over the whole period) to have zero mean and standard deviation equal to 1. The three distributions are statistically different at 1% confidence level.

Figure 2 shows the distribution of firm-level TFP over selected years, respectively 1995, 2000 and 2005. Over a decade, our estimations show that the productivity of French firms has increased significantly. The average manufacturing firm in 2005 is in fact about 8.6% percent more productive than the 1995 counterpart (the difference in mean of the two distributions is statistically significant at 1% level). The right shift of the distribution during the years suggests then a not negligible redistribution of firms towards higher levels of productivity. This does not mean however that the use of resources is increasingly close to optimal efficiency. Notwithstanding the developments in productivity, inefficiencies in factor allocation might be an obstacle to fully reap gains associated with technical progress. Looking at the within-industry productivity dispersion reveals a more heterogeneous picture, in 1995 the 90th to 10th percentile

[22] Results are reported in Appendix 6.2

ratio for French manufacturing firms was 1.04, meaning that for a given amount of inputs, most efficient firms were able to reach a level of production 189% more than low productive ones[23]. In 2007, given the average increase in productivity, the interquartile ratios increase to 1.17, suggesting that aggregate improvements were not driven by reallocation. It is worth noting that dispersion based on revenue productivity is usually smaller than the one computed on quantity-based productivity (see (Foster et al., 2008)), the reported values are likely to represent a lower bound for the true sectoral variability.

3. Resource (mis)allocation: aggregate and sectoral perspective

We now implement the method described in the previous Section and present the results obtained at the aggregate and sectoral level.

3.1. Sectoral gaps

The marginal productivity of production inputs is reported in the first two columns of Table 1. At the sectoral level input elasticity is always positive and very precisely estimated. Labor represents the highest coefficient in all industries as the input cost share. Estimated returns to scale are generally below unity and decreasing returns are indeed a sufficient condition for an optimal input choice without adjustment. However, we suspect imprecise measurement of certain inputs in sectors like pharmaceutical, a problem that would be fixed by redefining our production function with services as an input. This alternative specification, presented in Appendix, does not qualitatively change our results, although the distributive shares (and hence the labor gap) are slightly different (leading to slightly larger gaps).

Once the marginal productivity coefficients have been estimated the computation of the resource allocation gap is straightforward – from Equations (3) and (4). The main results for labor return to cost wedge are reported in Table 1.

[23]Since productivity is measured in log scale the percentage increase is given by $exp(1.04) - 1 = 189$.

Table 1 – Average Absolute Labour Gap by sector – years 1993-2007

Industry	Input Coefficients		Gap^{Abs}				Number
	β^l	β^k	Mean	CV	$Ineff^{Abs}_{\%}$	$Pos_{\%}$	Obs
Basic metals	0.576	0.242	10.936	0.787	36	15.5	10,274
Beverages	0.619	0.424	16.693	1.000	42	59.9	11,327
Chemicals	0.595	0.251	13.163	0.946	80	30.5	25,145
Computer and Elect	0.537	0.210	14.293	0.671	17	9.8	32,148
Electrical Equip	0.581	0.228	10.989	0.734	26	12.6	21,261
Fabricated metal	0.649	0.242	8.632	0.821	40	17.9	171,215
Food products	0.623	0.300	7.497	0.973	70	28.2	159,495
Furniture	0.609	0.232	8.984	0.719	21	10.9	35,939
Leather products	0.730	0.373	6.722	1.215	71	26.9	10,779
Machinery and Equip	0.646	0.192	10.198	0.850	44	19.4	65,093
Motor vehicles	0.650	0.284	8.670	0.956	52	19.7	17,050
Non-metallic pro	0.578	0.267	10.647	0.860	57	20.2	39,482
Other Manuf	0.624	0.314	10.105	0.865	53	22.0	45,095
Other transport	0.652	0.260	9.175	0.937	51	21.8	7,398
Paper products	0.641	0.245	9.034	0.930	61	26.0	18,602
Pharmaceutical	0.418	0.304	19.781	0.678	46	17.4	5,439
Printing and rec	0.661	0.171	9.457	0.853	37	17.3	79,861
Repair and instal	0.693	0.158	8.491	0.923	52	20.9	88,871
Rubber and plastic	0.609	0.204	8.979	0.889	54	22.3	45,964
Textiles	0.652	0.277	8.474	1.024	64	24.5	30,959
Wearing apparel	0.688	0.325	8.679	1.109	73	24.8	40,916
Wood products	0.644	0.249	6.922	0.952	54	22.9	43,673

For a given sector, k, the mean absolute Labor Gap is defined as follows, $\overline{Gap}_k^{Abs} = \frac{\sum_{i \in k} |G_i|}{N_k}$; it measures the distance from the social optimum allocation[24] where each firm is operating under perfect competition: i.e. marginal revenue equal to marginal costs and there are no frictions in the input markets. For the whole manufacturing sector over the period 1993-2007 this figure is slightly above 10 thousand euro per firm[25], but dispersion is relatively high not only between but also within industries, as shown by the coefficient of variation (CV). Instead of using the perfect competition (zero gap) as benchmark one may be interested to know what would be the contribution to overall gains in reaching the efficient allocation, i.e. all existing gaps are equal across firms in a given sector. This would result from lightening market constraints: allowing the reallocation of one unit of labor (i.e. the marginal worker) across firms without changing the employment level and the structural frictions. Such information is captured by the term $\text{Ineff}_{\%}^{Abs}$, derived as the ratio: $\frac{\sum_{i \in k} G_i/N_k}{Gap_k^{Abs}}$. In case of Electrical Equipment, for instance, structural frictions accounts for 26% of the mean absolute gap, while resource allocation inefficiencies determine the remaining 74%.

Looking closely at the last year of the sample we observe that about one fifth of manufacturing firms in the sample reports a positive wedge between labor marginal return and cost (see Table 2[26]).

Table 2 – Labor Gap decomposition, year 2007

	$\lvert\tilde{G}_{it}^l\rvert$	$G_{it}^l > 0$	$G_{it}^l < 0$
# of Firms	63,122	14,581	48,541
Share (%)	100	23	77
Mean	9.943	12.498	9.176
sd	9.247	14.351	6.840
10%	1.845	0.914	2.359
Median	7.724	6.930	7.845
90%	19.396	34.324	16.996

The sign of the gap is meaningful since it helps disentangling the variability and the direction

[24] Under the implicit assumption that the marginal worker has a productivity in line with the average of the observed firm.

[25] All the monetary values are expressed in real terms (euros of 2005), deflated using consumer price index.

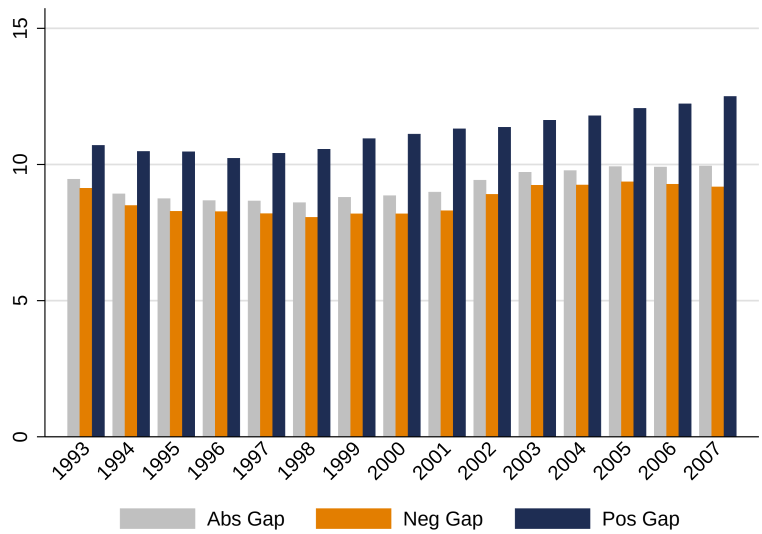
[26] The share of firms with positive gaps on the whole period is 22%.

of firm level mis-allocation. The average positive gap is roughly 36% higher than the negative counterpart, and the overall distribution for positive wedges seems to be relatively more right-skewed with respect to the negative ones. Assuming an average labor cost[27] in France of 50 thousand euro per year in 2007, an average negative wedge of 9.2 thousand euro implies that marginal return of labor is smaller than its cost of about 2.2 month salary. Notice that there is huge dispersion of such gaps and that the median negative wedge is at most of 7 thousand euros. On the other hand the value produced by the marginal worker is higher than its cost by almost 12.5 thousand euro when positive wedge are observed, although a firm should demand labor until its marginal return equals its costs under perfect competition.

3.2. Aggregate evolution of gaps

Over time the average gap seems to be slightly increasing after 2001 as reported in Figure 3, even without any control on firms' characteristics.

Figure 3 – Labor Gap, thousand Euros – real terms – manufacturing sector (unconditional mean)

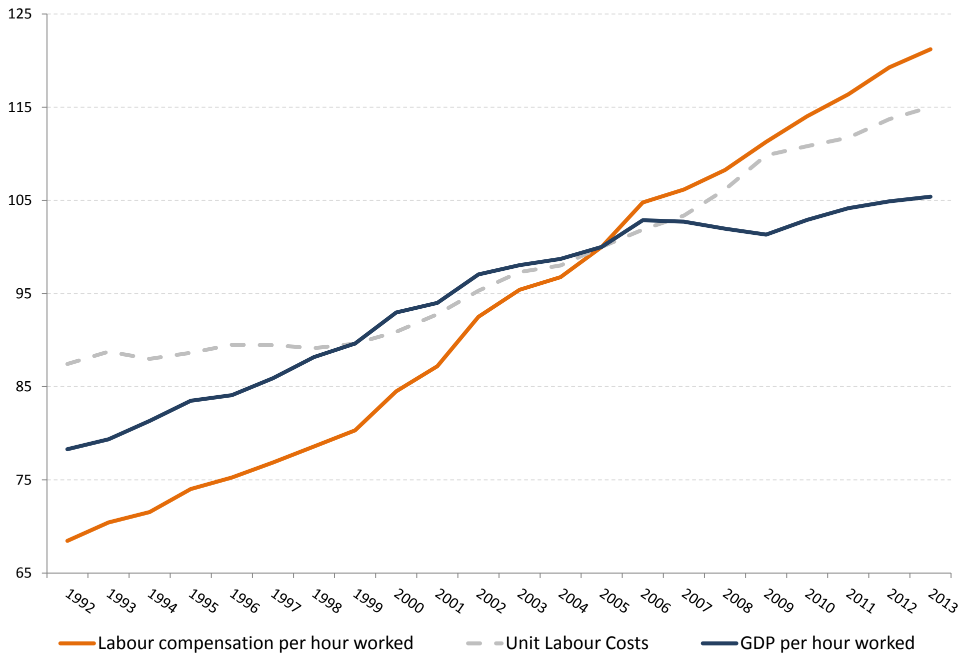


Note: Absolute Labor gap, unconditional mean.

[27] Including both salary and tax allowance.

From an aggregate point of view, Figure 4 shows that the French economy has experienced an increasing difference between marginal returns of labor and production costs during the 2000s. Aggregate Unit Labor Cost (ULC), often used as an indicator of country competitiveness, in fact, does not change significantly during the 1990s, afterwards the two underlying components start diverging and the ULC deteriorates: labor costs increase much faster than marginal returns. Given this aggregate stylized fact we should observe a similar behavior for the wedge between marginal productivity and marginal cost at the firm level.

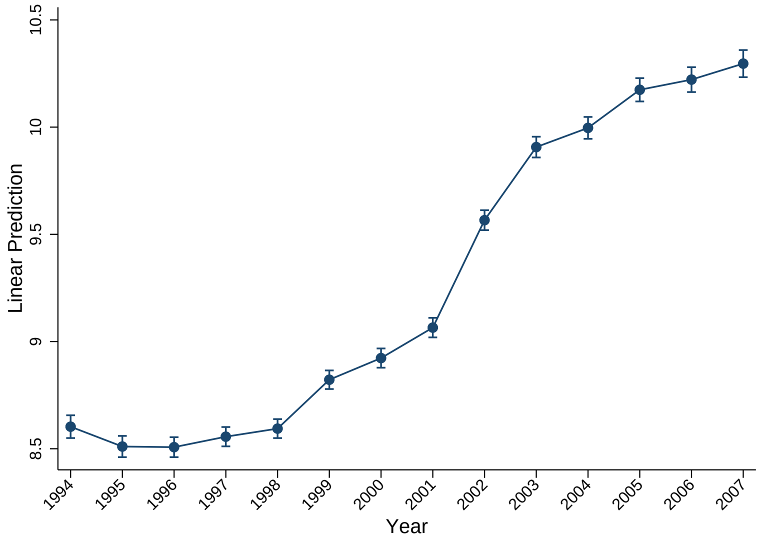
Figure 4 – Decomposition of the Unit Labor Cost



Source: OECD Productivity and ULC Dataset.

Figure 5 plots the monetary value of the average labor gap once we control for firms characteristics (fixed effects). The evolution over time of the firm level labor gap is highly similar to the aggregate figures on ULC, recording a significant jump after the 2001, suggesting that the economy has moved away from the zero gap benchmark (i.e. optimal allocation of resources). Interestingly, over the same period also the dispersion of labor gaps rose substantially, see Figure 7 in Appendix 6.1, indicating an associated increase of allocative inefficiencies. The main advantage of this methodology is twofold. First, it represents a useful tool in order to highlight firm level heterogeneity on the balance between marginal returns and costs. Second, it can help to identify the economic drivers of resource allocation efficiency at different aggregation levels.

Figure 5 – Absolute Labor Gap conditional on firm characteristics (average)



Note: Absolute Labor Gap by year keeping all covariates from the baseline estimation at their mean.

Labor gaps have indeed a geographic dimension. This mirrors to some extent differences in the productivity of the considered activity and skill-level of the mean wage in the observed firms. Thus one should not interpret the cross sectional evidence without controlling for the (here) unobserved characteristics of the location. In Section 4.2 we investigate further this issue by controlling for agglomeration economies, that may drive both wages and TFP[28] at the "Département" level.

4. Firm Level Evidence

In what follows we estimate the dynamics of the labor gap controlling for firm characteristics. Our aim is threefold. We firstly ask whether firms of different size face different obstacles to optimize their use of labour. If the external labour market is sticky, firms may resort to internal

[28]See for example **?**.

markets, and the more so for large firms resorting on a large internal pool of competencies. We then turn to our main question, i.e. whether firms in denser areas exhibit lower labour gaps, controlling for firm size and multi-plants. Finally, we confirm the presence of an additional channel of efficiency, the transmission of allocative inefficiencies downstream the value chains. The combination of the two latter sets of results provide a clear picture whereby denser areas provide better matching opportunities between employers and employees, controlling for firm size, while the proximity of suppliers also better optimizing reinforces the benefits of agglomeration.

4.1. Firm size

The baseline estimated equation is defined as:

$$Y_{it} = \alpha_0 + \delta_1 + \delta_2 + \delta_3 + \Gamma_{it}\beta + \xi_i + \upsilon_{it} \tag{5}$$

Where Y_{it} is the value of the absolute labor gap, $\left|G^l_{it}\right|$. The time evolution of the dependent variable is accounted by three sub-period dummies: δ_1 for the years 1998-2000, δ_2 for 2001-2003 and δ_3 for the last period (2004-2007). The constant α_0 captures the reference period gap value. The vector Γ_{it} includes a set of controls by firm and industry, namely a series of dummies identifying the quintile of production value by sector and year and an index for the degree of competition at the industry level, $Comp_{kt}$, computed as the $ln(1/HH)_{kt}$, where HH is an Herfindahl-Hirschmann index of employment concentration by sector k and year t. Finally, ξ_i are firm fixed effects to control for unobserved heterogeneity and υ_{it} is an idiosyncratic shock.

Main results of our analysis on the evolution of the labor gap for manufacturing firms are reported in Table 3. Controlling for firm fixed effects shows that for the average firm the wedge between marginal return and marginal cost of labor has increased significantly over the period, especially in the last years of the sample.

After slightly decrease in the first period, the average labor gap has increased significantly over time. In the period 2004-2007, in fact, the average gap is around 15 percent higher with respect to the reference period (column 2). Adding further controls on firms' size distribution (column 3) does not alter the main evidence. Interestingly quintile dummies shows that bigger firms tend to exhibit lower gaps, consistently with the empirical evidence showing that larger firms tend to be more productive.

Moreover, we find the same dynamics also when labor gap is conditioned on the transmitted component of productivity ω_{it}, column 1. The robustness check using the transmitted (and predictable) component of productivity is particularly relevant since the unpredictable component of the error term, ε_{it}, may come from measurement errors, and because many TFP estimations methods[29] assume that the firm choice of variable inputs is made after observing the productivity shock ω_{it}[30].

In Table 4 we perform a series of robustness check on the sensitivity of our results to sample selection. The evolution over time of firm input misallocations is consistent if we restrict the sample to firms with at least 20 employees ("restricted sample") or at small firms with less than 20 workers ("small firms"). Moreover, we find consistent evidence also if we restrict the sample only to single plant firms, signalling that our results should not be driven by compositional effects or by measurement errors induced by consolidated financial accounts (in case of multi-plant firms). For young firms (less than 5 year of activity) the increase in labor gap is milder but still significant[31].

The evolution over time of the labor gap for an average manufacturing firm with positive or negative values is shown in Figure 6, and the evolution for firms with less (more) than 49 employees is shown in Figure 10[32].

[29] Notably two step semi-parametric algorithms, like Levinsohn and Petrin, and related parametric estimators (GMM).

[30] A profit–maximizing firm should, then, equate the marginal product to the input cost, conditional on productivity ω_{it}.

[31] The results in Table 3 are robust to the use of the transmitted productivity ω_{it} to weight firm gaps.

[32] The two graphs reports the value of time dummies (interacted with an indicator variable for the characteristic of interest) holding firm characteristics' at their mean value – Equation (5).

Table 3 – Evolution of Labor Gap by selected period, real euro (thousand)

Dep. Var. :	Labor Gap $\lvert G^{l}_{it}\rvert$		
	(1)	(2)	(3)
1993-1997^{RefPeriod}	8.154***	8.890***	10.617***
	(0.900)	(0.266)	(0.361)
1998-2000	-0.257***	-0.115**	-0.125**
	(0.081)	(0.048)	(0.047)
2001-2003	0.909***	0.730***	0.714***
	(0.287)	(0.119)	(0.134)
2004-2007	3.072***	1.462***	1.468***
	(0.711)	(0.171)	(0.195)
$Comp_{kt}$	0.176	-0.039	0.014
	(0.172)	(0.057)	(0.067)
Size: 2^{nd} quintile			-1.125***
			(0.149)
Size: 3^{rd} quintile			-1.991***
			(0.237)
Size: 4^{th} quintile			-2.760***
			(0.308)
Size: 5^{th} quintile			-3.496***
			(0.416)
$\lvert G^{l}_{it}\rvert$ Wgt by ω_{it}	yes	no	no
Observations	1,005,986	1,005,986	1,005,986
R-squared	0.673	0.603	0.604

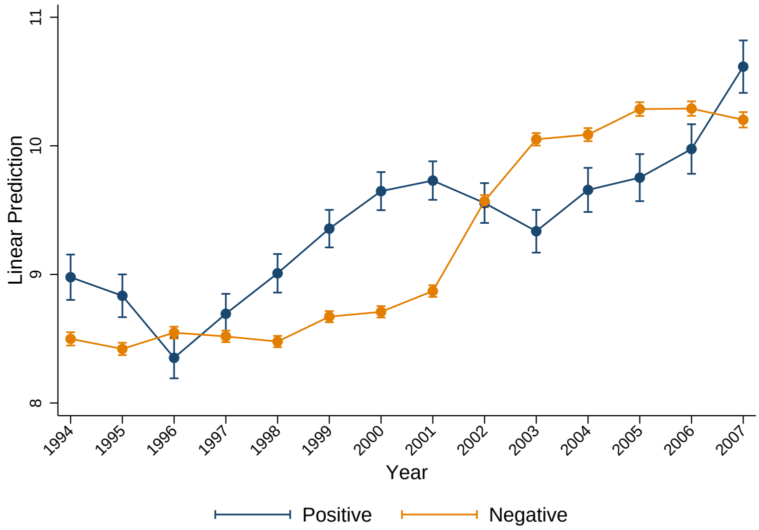
Standard errors clustered by industry k. All regressions include firm fixed effects. Quintile Size dummies are computed on sales distribution by sector and year. Dependent variable: labor gap in real euro. Marginal Productivity of labour is computed using Wooldridge (2009) modification of Levinsohn-Petrin algorithm, considering Capital as fixed input, Labor as flexible inputs and Raw Materials as proxy. Results in column (1) weights Gaps using firm level productivity. Results are robust to TFP estimation considering labor as fixed input.

Table 4 – Evolution of Labor Gap sample sensitivity

Dep. Var :	Labor Gap $\lvert G^{l}_{it}\rvert$			
	Restricted sample	Small firms	Young firms	Single Plant
1993-1997^{RefPeriod}	8.573***	8.760***	10.896***	10.256***
	(0.376)	(0.362)	(0.452)	(0.363)
1998-2000	0.070	-0.199***	-0.160***	-0.139***
	(0.068)	(0.044)	(0.050)	(0.049)
2001-2003	0.864***	0.700***	0.135	0.693***
	(0.112)	(0.137)	(0.161)	(0.126)
2004-2007	1.771***	1.372***	0.322	1.418***
	(0.183)	(0.172)	(0.276)	(0.181)
$Comp_{kt}$	-0.076	0.021	-0.095	0.008
	(0.073)	(0.069)	(0.083)	(0.072)
Size: 2^{nd} quintile			-0.726***	-1.092***
			(0.133)	(0.150)
Size: 3^{rd} quintile			-1.330***	-1.914***
			(0.185)	(0.228)
Size: 4^{th} quintile			-1.486***	-2.590***
			(0.295)	(0.297)
Size: 5^{th} quintile			-1.667***	-3.144***
			(0.481)	(0.443)
Observations	310,993	694,993	159,373	787,338
R-squared	0.645	0.613	0.751	0.602

Standard errors clustered by industry k. All regressions include firm fixed effects. Quintile Size dummies are computed on sales distribution by sector and year. Restrict $>$ 20 workers, Small $<$20 workers, Young $<$ 5 years of activity, Single Plant. Dependent variable: labor gap in real euro, computed using Wooldridge (2009) modification of Levinsohn-Petrin algorithm, considering Capital as fixed input, Labor as flexible inputs and Raw Materials as proxy.

Figure 6 – Average Labor Gap conditional on firm characteristics



The dynamics for negative and positive gaps is different. A sharp increase in the average negative gap is observed from 2001 on. This has a large impact on the average absolute gap, given the high frequency of negative gaps in the sample. This change is contemporary of new regulations on the labor market, but we cannot assess the causality. In contrast, the positive gap increased from the mid-nineties, notwithstanding a transitory stabilization in the middle of the period considered. Beyond identifying these two different evolutions, the method we use authorizes to disentangle two possible categories of determinants of the observed gaps. For negative values, the lack of optimization can be driven by distortions hampering the adjustment of firm to new market conditions. Different explanations can be considered for positive gaps, whereby firms can be kept below their optimal size – as defined under perfect competition – due to market imperfections such as market power. Our results are robust to the restriction of labor gaps to negative values. Thus, the potential drawback related to our assumptions on competition does not drive our conclusions.

As common to many countries labor regulation is more binding for bigger firms. In France this increasing stringent regulation is particularly relevant with more than 50 workers. From this threshold, in fact, firms must organize a works council; establish a committee for working conditions (health and safety); appoint a union representative[33]. The main effect of this increasing regulation is an increase in labor cost which may induce resource misallocation (see Garicano et al. (2013)) and potentially affect our results. From Figure 10 it emerges that firms choosing to stay below the 50 workers threshold report, other things equal, only a slightly higher gap, most likely due to the fact that those firms are operating at a sub-optimal scale. Despite this small difference the sharp increase in the wedge between labor marginal return and cost seems to have affected manufacturing firms irrespective to such threshold, confirming that our results may not driven by this distortion.

4.2. Agglomeration economies

We now come to our central argument and enrich the baseline specification in Equation 5 testing for the effect of agglomeration economies on return-cost wedges. Comparing the empirical firm productivity distribution across high and low density locations Combes et al. (2012) show that there is a substantial efficiency premium associated with city size, but it is even higher for highly productive firms. Interestingly, such premium is not related to selection but driven by agglomeration economies. Combes et al. (2012) are able to distinguish selection from agglomeration externalities thanks to a novel quantile approach that allows a close comparison of productivity distributions. Intuitively, this methodology relates the quantile of (log) productivity distribution in large and small cities to three key parameters: truncation, relative shift and dilation. A standard predictions of firm heterogeneity models is, in fact, that low productive firms should not survive in larger markets due to the higher degree of competition: productivity distributions should then display a left truncation in denser areas. However, Combes et al. (2012) do not find any evidence of left truncation (selection), instead, denser areas productivity distributions appear to be right shifted (average productivity premium) and dilated (more productive firms benefit more). Where the latter two characteristics are the results of the already mentioned agglomeration externality mechanisms: sharing, learning and matching.

[33] Above this threshold firms are also expected to establish a "plan social" when more than 9 employees are laid off at the same time (proving that the entrepreneur has been looking for another position for the dismissed).

In what follows we focus on the matching channel and test if in denser areas the thicker labor market also affect the firm resource allocation efficiency, i.e. return to cost wedge. Aiming to control for intra and inter industry agglomeration externalities we add to the vector Γ_{it} a set of measures on the economic environment at the "Département" level (NUTS3 administrative entities). In defining the indicators we follow Martin et al. (2011): for a firm i located in the Départment d and operating in the sector k we include:

- $Urbanization = ln\left(employees_t^d - employees_t^{dk} + 1\right)$
- $Location = ln\left(employees_t^{dk} - employees_{it}^{dk} + 1\right)$

where $Urbanization$ reports the number of employees in other industries within the same Départment d of industry i. This variable is meant to capture the inter-industry externalities, measured as the size of other industries' employment. $Location$, on the other hand, refers to intra-industry externalities measuring the number of employees working in the same industry k and the same Départment d as firm i[34]. In order to limit measurement errors induced by multi-plant firms the sample is restricted to single-plant firms (roughly 80% of the sample). We further exclude firms whenever they change location or sector during the estimation period (on average around 2.7% of the sample).

Results are reported in Table 5 and largely confirm previous findings about the timing of the increase of labor gaps. More interestingly, it is worth noting that, on average, *denser areas register lower gaps* across all specifications (in Column 1 and 2 $\left|G_{it}^l\right|$ is weighted by firm productivity). In terms of magnitude a 10% increase in the degree of urbanization is associated with a decrease in the average gap of roughly 196 euro, that this is equivalent to 21% of Labor Gap standard deviation (9,299 euro). On the other hand, highly specialized Departments do not seem to register a sizeable difference in the value of the labor gap, moreover the estimated coefficient is not statistically different from zero when the gap is not weighted by ω_{it} - columns 3 and 4.

[34]Noteworthy, from a firm point of view, the two measures along with its own number of employees describe exhaustively local employment in manufactures.

Table 5 – Agglomeration Externalities: Baseline Results

Dep. Var.:	Labor Gap $\lvert G^{l}_{it}\rvert$			
	(1)	(2)	(3)	(4)
1993-1997$^{Ref Period}$	27.853***	28.718***	16.377***	16.510***
	(2.881)	(2.828)	(1.025)	(1.040)
1998-2000	-0.167***	-0.160***	-0.102***	-0.101***
	(0.029)	(0.029)	(0.027)	(0.027)
2001-2003	0.924***	0.919***	0.667***	0.667***
	(0.046)	(0.046)	(0.035)	(0.035)
2004-2007	2.740***	2.718***	1.311***	1.308***
	(0.094)	(0.094)	(0.048)	(0.048)
$Urbanization_{dkt}$	-1.962***	-1.843***	-0.635***	-0.617***
	(0.290)	(0.287)	(0.104)	(0.104)
$Location_{idkt}$		-0.273***		-0.042
		(0.079)		(0.049)
$Comp_{dkt}$	-0.236**	-0.275**	-0.033	-0.039
	(0.117)	(0.117)	(0.042)	(0.043)
$\lvert G^{l}_{it}\rvert$ Wgt by ω_{it}	Yes	Yes	No	No
Observations	765,479	765,479	765,479	765,479
R-squared	0.681	0.681	0.607	0.607
Model	Level-Log	Level-Log	Level-Log	Level-Log

Standard errors clustered by *département* and industry dk. All regressions include firm fixed effects, time and size quintile dummies (not reported). Single plant firms only, excluding either *département* or sector relocations. $Comp_{dkt}$ is computed by sector k, *département* d and year t. Results are robust using $Comp_{kt}$.

In Table 6 we control for the possible correlation between our agglomeration externality variables and the error term; local economic shocks, in fact, may induce firms to change their employment profile inducing a simultaneity bias in the estimated coefficients. In order to control for this source of bias we adopt the following strategy: we take first differences of our baseline equation to remove the additive firm specific fixed effect and then estimate the transformed model using firm fixed effects. This procedure ensures that also the firm specific time trend is conditioned out (i.e. random trend models)[35]. We start from the following version of our baseline equation (Wooldridge, 2010):

$$Y_{it} = \xi_i + g_i t + \lambda_t + \Gamma_{it}\beta + \nu_{it} \tag{6}$$

Where ξ_i represents the individual time invariant heterogeneity, while $g_i t$ is an individual trend accounting for unobserved (time varying) heterogeneity, taking first difference of equation (6) we obtain:

$$\Delta Y_{it} = g_i + \rho_t + \Delta\Gamma_{it}\beta + \delta\nu_{it} \tag{7}$$

Equation (7) can now be estimated using Fixed Effects (or first-differencing again) to condition out individual specific trend g_i as well. Note that if Γ_{it} contains a time trend this becomes a constant after first differencing the data and it is conditioned out as well.[36]. Intuitively, removing most of the heterogeneity from Γ_{it} will make more likely to get consistent estimation.

Results, reported in Table 6, largely confirm previous findings. The degree of urbanization at the Département level is associated with a lower average labor gap, while the *Location* (i.e. local sectoral specialization) seems to only marginally affect labor input allocation. One explanation

[35]This class of models was proposed originally by Heckman and Hotz (1989), while Wooldridge (2005) provide a theoretical discussion.

[36]Note that the choice between fixed effects or first difference in estimating Equation (7) depends on the properties of $\delta\nu_{it}$ if it shows a large amount of serial correlation first difference may be more appropriate. In our case regressing the residuals from Equation (6) on their lagged value, gives a coefficient of -0.37 (se .002), which is different from zero but still different from -0.5 expected if there were no correlation in the original Equation 5; we choose to estimate Equation (7) using fixed effects. Results are robust if we use first difference instead.

may be related to the effect of idiosyncratic shocks: for highly specialized economies it may be more difficult to efficiently reallocate inputs after a demand shock with respect to more diversified ones. Regarding the magnitude (column 1) it is worth noting that now the effect of *Urbanization* is significantly lower, at least when labor gap is weighted by firm productivity shocks (-0.690). Such upward bias of the fixed effects estimation is not surprising if individual trends are correlated with local characteristics: i. e. positive correlation between firm productivity and *urbanization*. Once firms specific trends in ω_{it} are accounted for the negative correlation between agglomeration economies and labor gaps is marginally lower but still negative and highly significant.

In general, our results on labor misallocation gaps and urbanization suggests that part of the productivity premium for denser areas may be also due to a better allocation of resources (labor) across firms, confirming that one of the mechanism behind the productivity advantage of cities, Combes et al. (2012), is the mentioned "matching" channel.

4.3. Input-Output linkages

Growing empirical and theoretical evidence has shown that micro economic shocks may propagate throughout firms' interconnections and generate aggregate volatility, i.e. in production output (Acemoglu et al., 2012) or exports sales (Di Giovanni et al., 2014). Providing that firms are usually embedded in a complex production network any shock affecting suppliers' prices may propagate downstream through input-output linkages and affect the overall state of the economy (Acemoglu et al., 2015).

In what follows we test if there is evidence of positive correlation of firm level inefficiencies across sectors. Unfortunately information on firm specific linkages is not available so we rely on industry Input-Output (IO) table provided by the INSEE assuming that i) within sectors intermediate shares are homogenous and ii) there is a representative input provider for all firms.[37]

[37] We borrow these assumptions from Di Giovanni et al. (2014).

Table 6 – Agglomeration Externalities: Random Trend Specification

Dep. Var.:	Labor Gap $\lvert G^l_{it}\rvert$			
	(1)	(2)	(3)	(4)
1993-1997$^{Ref.Period}$	0.081***	0.012***	-0.007	0.005
	(0.015)	(0.002)	(0.020)	(0.004)
1998-2000	0.026	-0.000	0.067*	0.002
	(0.021)	(0.003)	(0.037)	(0.007)
2001-2003	0.263***	0.029***	0.315***	0.039***
	(0.019)	(0.003)	(0.032)	(0.004)
2004-2007	0.388***	0.035***	0.252***	0.014
	(0.034)	(0.004)	(0.050)	(0.010)
$Urbanization_{dkt}$	-0.672***	-0.059***	-0.423***	-0.050**
	(0.198)	(0.019)	(0.147)	(0.020)
$Location_{idkt}$	-0.105**	-0.008	-0.017	-0.002
	(0.045)	(0.007)	(0.051)	(0.008)
$Comp_{dkt}$	-0.178***	-0.022***	-0.075	-0.011
	(0.066)	(0.007)	(0.077)	(0.009)
$\lvert G^l_{it}\rvert$ Wgt by ω_{it}	Yes	Yes	No	No
Observations	618,330	618,330	618,330	618,330
R-squared	0.109	0.073	0.108	0.072
Model	Level-Log	Log-Log	Level-Log	Log-Log

Standard errors clustered by *département* and industry dk. All regressions include firm fixed effects, time and size quintile dummies (not reported). Single plant firms only, excluding either *département* or sector relocations. $Comp_{dkt}$ is computed by sector k, *département* d and year t. Results are robust using $Comp_{kt}$.

Our preferred specification makes use of the IO table from 1995 (first available year)[38] and considers for each firm i the average absolute gap in the most important supply (1^{st} Upstream) and demand (1^{st} Downstream) sector[39]. Results are reported in Table 7 where the coefficients are estimated using a random trend model to control for firm level unobserved heterogeneity.

In the bottom part of Table 7 we consider the average misallocation across all supply and demand sectors of firm i, each of them weighted using the share on sector i's total intermediate (indirect) requirements. Results point to a positive and significant cross-sectoral correlation of misallocation Gaps. Consistently with previous results this correlation is much stronger for

[38] Results are robust to the use of different release of IO tables, Table ST2 uses intermediates' shares from 2000 IO tables.
[39] Excluding the share of intermediates bought from (sold to) the same sector, i.e. direct requirements.

supply side relationship when Gaps are weighted by firms' productivity (ω_{it}). Notice that such correlation does not necessarily mean transmission of allocative inefficiency from one firm to another. Results are robust when Gaps do not take into account ω_{it} – column 3 and 4 – even if in this case supply and demand correlations are not statistically different.

Interestingly, such results confirm that firm inefficiencies are likely to diffuse throughout input-output linkages, the significant correlation between upstream and downstream efficiency levels indicates that the potential gain from improving inputs allocation across firms would be magnified by such externalities.

Table 7 – Mis-Allocation through Production Network

Dep. Var	Labour Gap $\|G\|^{l}_{it}$			
	(1)	(2)	(3)	(4)
	Random Trend Model			
Upstream Gap (1^{st} Supplier)	0.417***	0.511***	0.211***	0.291***
	(0.103)	(0.176)	(0.035)	(0.060)
Downstream Gap (1^{st} Buyer)	0.175**	0.230**	0.243***	0.311***
	(0.074)	(0.110)	(0.043)	(0.070)
$Comp_{kt}$	-0.197	-0.030	-0.159	-0.025
	(0.227)	(0.034)	(0.232)	(0.035)
Observations	618,330	618,330	618,330	618,330
R-squared	0.110	0.074	0.108	0.073
Upstream Gap (All Suppliers)	0.539***	0.684**	0.151**	0.275*
	(0.155)	(0.325)	(0.060)	(0.133)
Downstream Gap (All Buyers)	0.347***	0.465*	0.518***	0.617***
	(0.118)	(0.241)	(0.072)	(0.160)
$Comp_{kt}$	0.053	-0.006	-0.092	-0.023
	(0.214)	(0.040)	(0.128)	(0.023)
$\|G^{l}_{it}\|$ Wgt by ω_{it}	Yes	Yes	No	No
Observations	618,330	618,330	618,330	618,330
R-squared	0.110	0.074	0.108	0.073
Model	Level-Level	Log-Log	Level-Level	Log-Log

Standard errors clustered by industry k. All regressions include time and size quintile dummies (not reported). Results are obtained using a fixed effect estimator on first difference variables. Input-Output linkages refer to the first available year for IO tables, 1995.

5. Conclusion

Firms in denser areas are more productive. We argue that the gap between the value of the marginal product and marginal input price, which reveals inefficiencies in inputs allocation, is reduced in agglomerated locations. The nice feature of this approach using a reasoning at the margin, is to give a monetary value to this misallocation, and to disentangle positive and negative gaps. Using a methodology proposed by Petrin and Sivadasan (2013) we were able to assess the degree of resource misallocation at the firm level using French BRN data. The location of the firm (within French *Départements*) is observed, which informs on the degree of misallocation within sectors among locations of different density confronted to a common regulatory framework (e.g. labour market regulations). The average (marginal) gap at firm level over the period 1993-2007 is around 10 thousands euro.

We confirm that misallocation has a spatial dimension: resource allocation and the associated effect on productivity is not only related to firms characteristics but it is also related to the environment in which they operate. Denser locations offer a better match between employers and employees. Urbanization at the *Départment* level is associated with a lower average labor gap, suggesting that such matching is playing a role in determining the productivity advantage of denser areas. Furthermore, we also observe that firm inefficiencies are diffused throughout input-output linkages due to indirect inter-sectoral spillovers.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Networks, Shocks, and Systemic Risk. NBER Working Papers 20931, National Bureau of Economic Research, Inc.

Bartelsman, E., Haltiwanger, J., and Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *The American Economic Review*, 103(1):305–334.

Bartelsman, E. J., Haltiwanger, J. C., and Scarpetta, S. (2009). Cross-country differences in productivity: The role of allocation and selection. Working Paper 15490, National Bureau of Economic Research.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2007). Firms in international trade. *The Journal of Economic Perspectives*, 21(3):105–130.

Berthou, A. and Sandoz, C. (2014). Labour productivity in europe: allocative efficiency of labour or performance of firms? Bulletin Trimestriel 34, Banque de France.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D., and Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, 80(6):2543–2594.

Davis, S. J., Haltiwanger, J., Jarmin, R., and Miranda, J. (2007). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms. In *NBER Macroeconomics Annual 2006, Volume 21*, pages 107–180. MIT Press.

Dhingra, S. and Morrow, J. (2014). Monopolistic competition and optimum product diversity under firm heterogeneity. mimeo.

Di Giovanni, J., Levchenko, A. A., and Méjean, I. (2014). Firms, destinations, and aggregate fluctuations. *Econometrica*, 82(4):1303–1340.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In Henderson, J. V. and Thisse, J. F., editors, *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics*, chapter 48, pages 2063–2117. Elsevier.

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability? *American Economic Review*, 98(1):394–425.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.

Garicano, L., LeLarge, C., and Reenen, J. V. (2013). Firm size distortions and the productivity distribution: Evidence from france. Working Paper 18841, National Bureau of Economic Research.

Haltiwanger, J. (2011). *Innovation Policy and the Economy*, volume 12, chapter Job Creation and firm dynamics in the US, pages 17–38. University of Chicago Press.

Heckman, J. J. and Hotz, V. J. (1989). Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training. *Journal of the American statistical Association*, 84(408):862–874.

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448.

Lazear, E. P. and Spletzer, J. R. (2012). Hiring, churn and the business cycle. *American Economic Review: Papers and Proceedings*, 102(3):575–579.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341.

Martin, P., Mayer, T., and Mayneris, F. (2011). Spatial concentration and plant-level productivity in france. *Journal of Urban Economics*, 69(2):182–195.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica*, 64(6):1263–1297.

Petrin, A., Poi, B. P., and Levinsohn, J. (2004). Production function estimation in stata using inputs to control for unobservables. *Stata journal*, 4:113–123.

Petrin, A. and Sivadasan, J. (2013). Estimating lost output from allocative inefficiency, with an application to chile and firing costs. *Review of Economics and Statistics*, 95(1):286–301.

Syverson, C. (2004). Product substitutability and productivity dispersion. *Review of Economics and Statistics*, 86(2):534–550.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–65.

Syverson, C. (2014). The importance of measuring dispersion in firm-level outcomes. *IZA World of Labor*, 53(.):.

Van Biesebroeck, J. (2007). Robustness of productivity estimates*. *The Journal of Industrial Economics*, 55(3):529–569.

Wagner, J. (2012). International trade and firm performance: a survey of empirical studies since 2006. *Review of World Economics*, 148(2):235–267.

Wooldridge, J. M. (2005). Fixed-effects and related estimators for correlated random-coefficient and treatment-effect panel data models. *Review of Economics and Statistics*, 87(2):385–390.

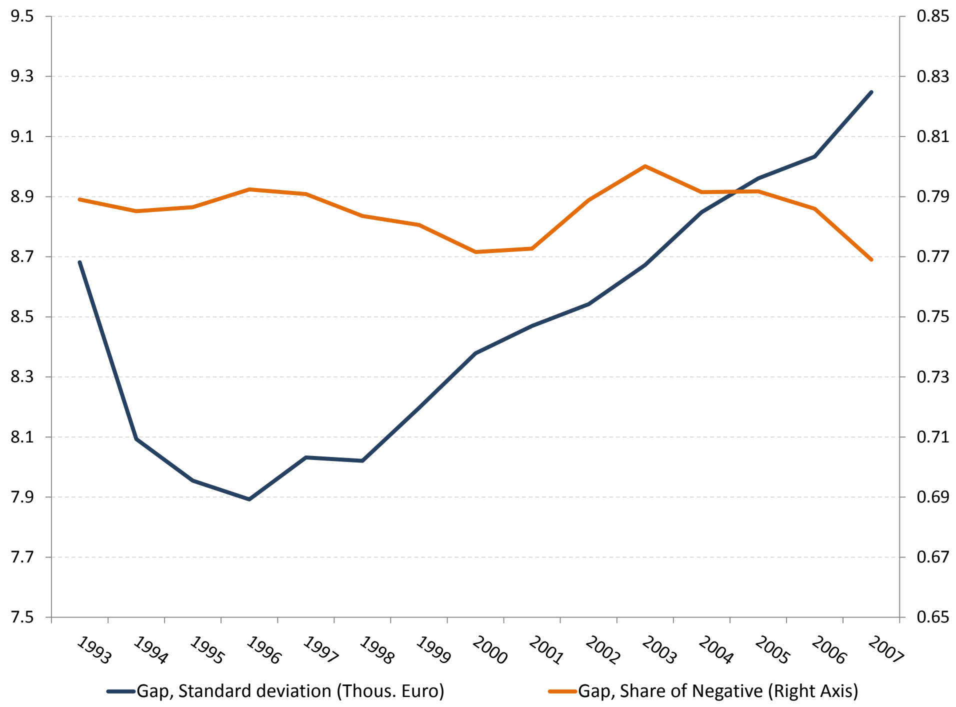
Wooldridge, J. M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Economics Letters*, 104(3):112–114.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

6. Appendix

6.1. Evolution of Labor Gap over Time & Sectors

Figure 7 – Absolute Labor Gap over Time: Dispersion and Composition



In the following we present additional evidence on the evolution of labor gap over time. Figure 7 reports the dispersion (standard deviation) of the labor gap; the sudden increase in the dispersion of firm gaps suggest an increase in the overall inefficiency of input allocation across firms, this effect does not seems to be driven by the composition of the gap: the share of negative in 2013, in facts, is only few percentage points below the 1994 values.

The sectoral evolution of Labor Gap is reported in Table ST1. Column (a) shows the average misallocation gap in the the reference period, from 1993 to 1997, where the degree of inefficiency was relatively stable; while column (b) refers to the period 2004-2007 both values are expressed in real (thousand) euros (of 2005), deflated using the Consumer Price Index. Over time the average gap increased significantly – on average of about 13% – but with relevant heterogeneity across sectors, e.g. Transport Equipment (26%), Wearing Apparel (21%) and Chemicals (20%).

In monetary terms the largest increase is in Beverages where the average gap in the last period is over 3 thousands euro higher with respect to the reference period. Interestingly, there is a large degree of heterogeneity also in the share of firms registering positive gaps, from around 19% in Basic metals production to over 60% for Beverages (column - c). Positive gaps are particularly interesting since they illustrate how many firms, within the sector, are operating at a sub-optimal scale (indeed, positive gaps may reflect market power – i. e. markups – in some highly differentiated sectors e.g. beverages). The industry specific value added gain from reallocation of a marginal worker in the optimal direction, i.e. from "lower to higher marginal value activities" (Petrin and Sivadasan, 2013) is shown in Figure 8. For France this value is equal to roughly 0.38% of manufacturing value added, a non-negligible figure considering it is computed at the margin.

Table ST1 – Average Absolute Labor Gap by sector over time

Industry	Mean Gap (Thous. Euro)		Positive Gap (%)		ΔGap	ΔPositive
	'93/'97	'04/'07	'93/'97	'04/'07		
	(a)	(b)	(c)	(d)	(b)-(a)	(d)-(c)
Basic metals	10.558	11.963	13.2	19.1	1.405	5.91
Beverages	14.917	18.175	56.3	60.7	3.259	4.46
Chemicals	12.030	14.517	29.5	31.3	2.487	1.81
Computer and Elect	13.834	15.311	8.6	11.2	1.477	2.58
Electrical Equip	10.574	11.769	11.1	13.6	1.195	2.49
Fabricated metal	8.349	9.164	16.0	19.5	0.815	3.55
Food products	7.480	7.700	31.7	23.5	0.220	-8.24
Furniture	8.861	9.674	10.3	11.2	0.813	0.88
Leather products	6.525	7.522	26.1	27.8	0.997	1.63
Machinery and Equip	9.870	10.872	17.6	21.3	1.002	3.72
Motor vehicles	8.367	9.599	18.6	20.6	1.233	1.95
Non-metallic pro	10.151	11.554	17.5	23.2	1.404	5.71
Other Manuf	9.752	11.047	22.1	21.0	1.294	-1.12
Other transport	8.336	10.568	18.8	22.6	2.232	3.83
Paper products	8.554	9.608	25.6	24.8	1.053	-0.79
Pharmaceutical	19.148	20.801	16.0	17.0	1.653	0.93
Printing and rec	9.159	10.066	17.2	17.3	0.907	0.12
Repair and instal	8.234	9.042	19.1	22.4	0.808	3.27
Rubber and plastic	8.607	9.629	21.1	22.6	1.022	1.59
Textiles	8.333	9.161	26.6	20.8	0.829	-5.85
Wearing apparel	8.268	9.998	24.0	25.5	1.730	1.51
Wood products	6.657	7.610	20.6	25.1	0.954	4.45

Figure 8 – Value Added Gain from reallocation, a unit of labor is moved from one establishment to another

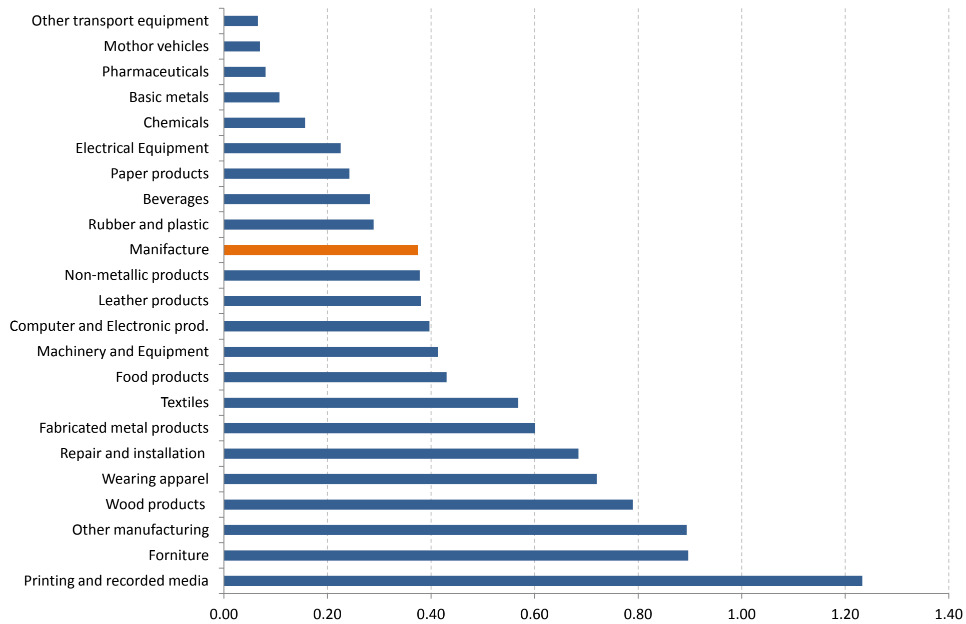


Figure 9, plots gaps for single-plant firms conditional on their location in high versus low density Départments, where the high density are those above the median value by year. This evidence further characterizes results shown in Table 6. On average, firms located in highly urbanized Départments observe lower gaps, but this advantage is weakening over time. More precisely, the significant difference observed at the beginning of the period that confirms the above mentioned hypothesis of a better matching in agglomerated economies, vanishes over the period, in line with the spread of larger labor gaps over the French map.

Another issue to tackle, is the relation between the labor gaps and the dynamics of firms' size. There is ample evidence that a discontinuity is present in the French demography of firms around the 50 employees threshold, the latter corresponding to specific regulations imposed to firms in the social arena. We confirm that such discontinuity is present in our data, but it results only in a marginal difference in labor gap levels not in their evolution over time (see Figure 10). Interestingly it appears that firms choosing to stay below the 50 workers threshold report, other things equal, a slightly higher gap; most likely due to the fact that they are operating at a sub-optimal scale. Despite this small difference the sharp increase in the wedge between labor marginal return and marginal cost seems to have affected manufacturing firms irrespective to such break.

Finally, Figure 11 largely confirms the evolution of the labor gap even when conditional on firm productivity (ω_{it}).

Figure 9 – Avg. (absolute) Labor Gap by Density

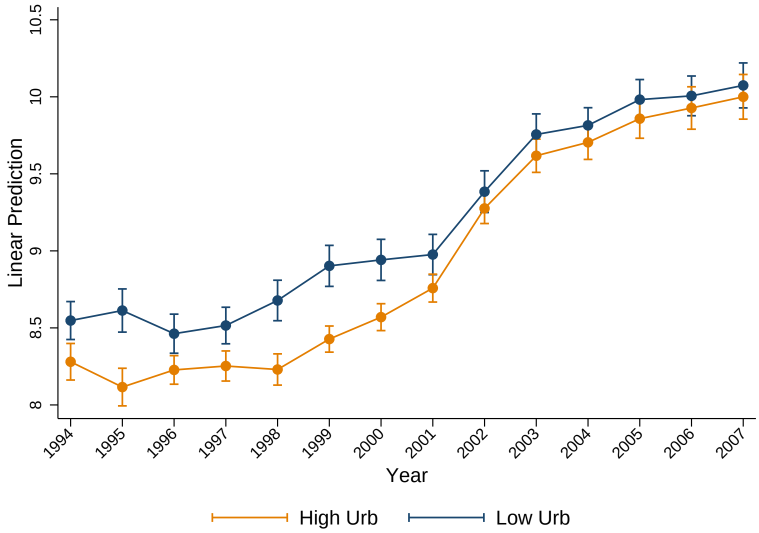


Figure 10 – Avg. (absolute) Labor Gap conditional on Firm Size (Number of Workers)

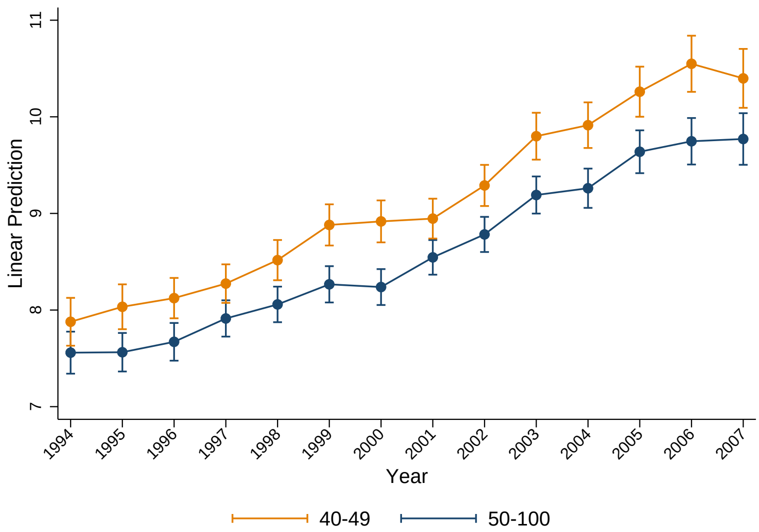


Figure 11 – Avg. (absolute) Labor Gap conditioned on ω_{it}

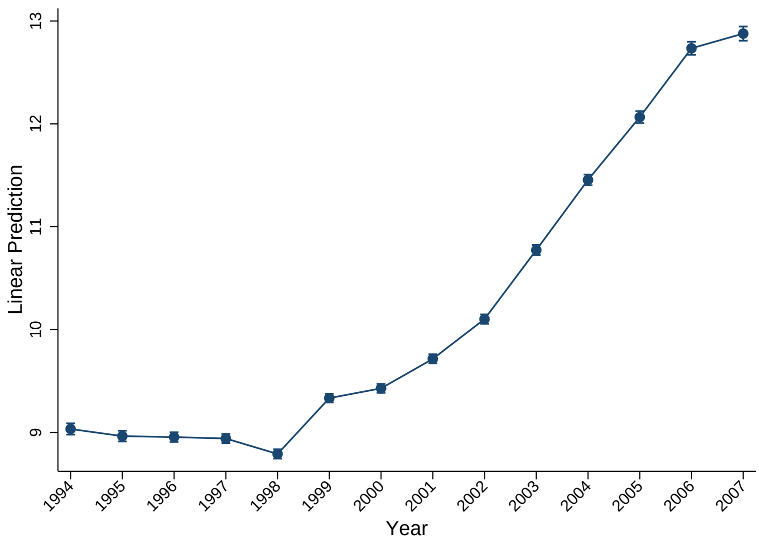


Table ST2 – Mis-Allocation through Production Network: IO 2000

Dep. Var	Labour Gap $\lvert G\rvert_{it}^{l}$			
	(1)	(2)	(3)	(4)
	Random Trend Model			
Upstream Gap (1^{kt} Supplier)	0.398***	0.494**	0.230***	0.308***
	(0.117)	(0.186)	(0.039)	(0.065)
Downstream Gap (1^{kt} Buyer)	0.213***	0.234***	0.248***	0.300***
	(0.065)	(0.079)	(0.040)	(0.055)
$Comp_{kt}$	-0.174	-0.027	-0.022	-0.007
	(0.226)	(0.035)	(0.188)	(0.031)
Observations	618,330	618,330	618,330	618,330
R-squared	0.110	0.074	0.108	0.073
Upstream Gap (All Suppliers)	0.551***	0.715**	0.159***	0.288**
	(0.154)	(0.326)	(0.056)	(0.129)
Downstream Gap (All Buyers)	0.322***	0.416*	0.503***	0.597***
	(0.113)	(0.234)	(0.071)	(0.155)
$Comp_{kt}$	0.028	-0.010	-0.078	-0.021
	(0.213)	(0.039)	(0.126)	(0.023)
$\lvert G_{it}^{l}\rvert$ Wgt by ω_{it}	Yes	Yes	No	No
Observations	618,330	618,330	618,330	618,330
R-squared	0.110	0.074	0.108	0.073
Model	Level-Level	Log-Log	Level-Level	Log-Log

Standard errors clustered by industry (k). All regressions include time and size quintile dummies (not reported). Results are obtained using a fixed effect estimator on first difference variables. Input-Output linkages are built on 2000 IO tables.

6.2. Robustness to Alternative TFP estimations

In the following section we report the results obtained using different approaches to the computation of the TFP and the implied marginal productivity of labor, building block of firms misallocation variable. The estimation of the firm specific labor gap here starts from a slightly different Cobb-Douglas production function for firm i at time t defined as the following:

$$q_{it} = \beta_l l_{it} + \beta_k k_{it} + \beta_s s_{it} + \omega_{it} + \varepsilon_{it} \tag{8}$$

Where q_{it} denotes value added, l_{it} the number of employees, k_{it} the fixed capital stock and s_{it} is the demand for services (including energy). Note that here value added is defined as revenues minus outside purchases, where the latter include only material inputs (m_{it}); meaning that we now are taking out services from the material inputs aggregate and considering them as a factor of production [40]. Using this strategy we are able to consider Labor, along with Capital, as a fixed production factor (i.e. assuming that there are friction in the labor markets) and using Services as freely adjustable inputs in the TFP estimation.

To ease comparability with previous findings we report the results obtained estimating Equation (8) assuming Labor as flexible input (as in the baseline reported in the text), see Table ST3[41] and considering Labor as fixed input[42], see Table ST4.

Previous findings commented in the core of the text are robust to such change in the specification of the production function: labor gap has significantly increase over the 1993-2007 period.

[40] As before, all the variable are in logs and deflated using industry price indexes from the INSEE.
[41] In such case Capital is instrumented using m_{it} while l_{it} and s_{it} are considered free to adjust to productivity shocks.
[42] Capital and Labor are instrumented using m_{it} while s_{it} is the only free input.

Table ST3 – Marginal Product (MP) Estimated from Eq. 8, Labor as Free Input

Industry	Labor	Capital	Services	RTS	P-Value for CRS
Basic metals	0.286	0.137	0.533	0.957	0.05
Beverages	0.283	0.202	0.619	1.105	0.00
Chemicals	0.279	0.097	0.619	0.995	0.76
Computer and Electr	0.313	0.096	0.515	0.924	0.00
Electrical Equip	0.321	0.114	0.521	0.957	0.01
Fabricated metal	0.368	0.126	0.489	0.982	0.00
Food products	0.353	0.158	0.490	1.000	0.95
Furniture	0.340	0.123	0.525	0.987	0.27
Leather products	0.442	0.168	0.518	1.128	0.00
Machinery and Equip	0.329	0.074	0.530	0.934	0.00
Motor vehicles	0.374	0.128	0.497	1.000	0.99
Non-metallic pro	0.300	0.110	0.542	0.952	0.00
Other Manuf.	0.387	0.187	0.458	1.032	0.02
Other transport	0.356	0.087	0.589	1.032	0.20
Paper products	0.322	0.127	0.527	0.976	0.13
Pharmaceutical	0.175	0.080	0.707	0.962	0.34
Printing and record	0.349	0.078	0.545	0.971	0.00
Repair and instal	0.389	0.069	0.494	0.952	0.00
Rubber and plastic	0.327	0.075	0.518	0.920	0.00
Textiles	0.352	0.102	0.540	0.993	0.63
Wearing apparel	0.420	0.143	0.546	1.108	0.00
Wood products	0.362	0.106	0.499	0.967	0.01

Note: Productivity Coefficients estimated using Wooldridge (2009) GMM approach on a value added production function, considering Capital as fixed and Labor and Electricity (plus Services) as flexible inputs and Raw Materials as proxy. In the last column the Null hypothesis is $\beta_l + \beta_k + \beta_s = 1$. All input coefficients are significant at 1% level - β_k for pharmaceutical at 5%.

Table ST4 – Marginal Product (MP) Estimated from Eq. 8, Labor as Fixed Input

Industry	Labor	Capital	Services	RTS	P-Value for CRS
Basic metals	0.333	0.128	0.513	0.974	0.41
Beverages	0.240	0.196	0.620	1.055	0.13
Chemicals	0.329	0.080	0.605	1.014	0.53
Computer and Electr	0.315	0.093	0.509	0.918	0.00
Electrical Equip	0.349	0.105	0.510	0.964	0.07
Fabricated metal	0.311	0.134	0.487	0.933	0.00
Food products	0.273	0.167	0.482	0.922	0.00
Furniture	0.311	0.128	0.518	0.957	0.00
Leather products	0.433	0.154	0.506	1.093	0.00
Machinery and Equip	0.311	0.075	0.528	0.914	0.00
Motor vehicles	0.380	0.128	0.480	0.988	0.58
Non-metallic pro	0.315	0.106	0.525	0.946	0.00
Other Manuf.	0.365	0.187	0.448	1.000	1.00
Other transport	0.329	0.095	0.582	1.006	0.86
Paper products	0.298	0.129	0.528	0.954	0.02
Pharmaceutical	0.314	0.023	0.684	1.021	0.65
Printing and record	0.269	0.088	0.545	0.901	0.00
Repair and instal	0.321	0.082	0.493	0.896	0.00
Rubber and plastic	0.332	0.072	0.512	0.916	0.00
Textiles	0.387	0.105	0.522	1.014	0.40
Wearing apparel	0.348	0.154	0.537	1.038	0.00
Wood products	0.327	0.106	0.489	0.922	0.00

Note: Productivity Coefficients estimated using Wooldridge (2009) GMM approach on a value added production function, considering Capital and Labor as fixed, Electricity (plus Services) as flexible input and Raw Materials as proxy. In the last column the Null hypothesis is $\beta_l + \beta_k + \beta_s = 1$. All input coefficients are significant at 1% level - β_k for pharmaceutical is not significant.

Figure 12 – Average Labor Gap: marginal product from Eq. 8, Labor as Free Input

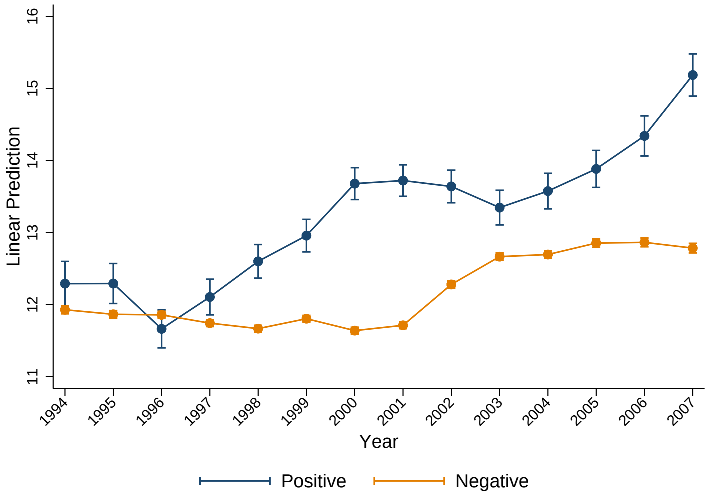
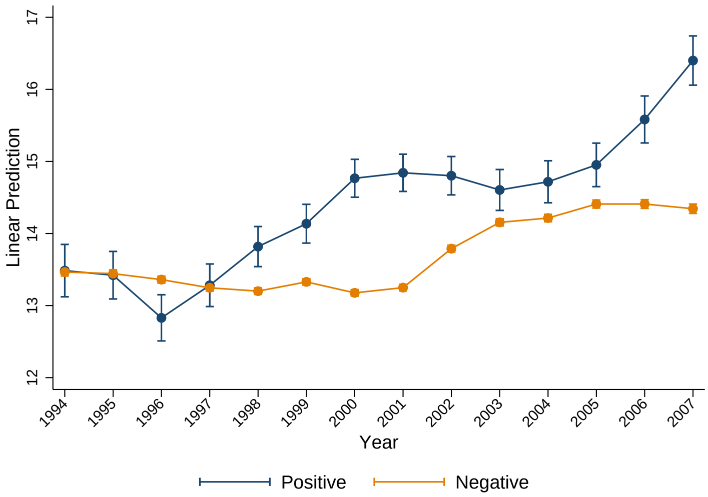


Figure 13 – Average Labor Gap: marginal product from Eq. 8, Labor as Fixed Input



6.3. Data

The dataset covers the period 1993-2007. We stop our exercise in 2007 in order to leave aside the contrasted reaction of firms to the subsequent economic crisis. After excluding implausible observations, namely those reporting negative or zero values for our variables of interest and cleaning the data from potential outliers[43], we end up with an un-balanced panel of 137,119 firms for the French manufacturing sector[44]. Single plant firms represent 80% of the observations, meaning that in the vast majority of the cases we observe production functions at the plant level.

Table ST5 – Number of firms in the estimation sample (by year)

Year	Firms	% Single Plant	% Within Same Dep.
1993	69,740	0.795	0.891
1994	68,268	0.805	0.897
1995	69,232	0.811	0.901
1996	67,728	0.811	0.900
1997	69,407	0.809	0.899
1998	68,849	0.807	0.897
1999	68,624	0.807	0.897
2000	67,798	0.801	0.895
2001	66,409	0.795	0.891
2002	67,241	0.791	0.890
2003	66,557	0.790	0.889
2004	65,717	0.790	0.889
2005	64,232	0.779	0.883
2006	63,062	0.777	0.882
2007	63,122	0.779	0.884

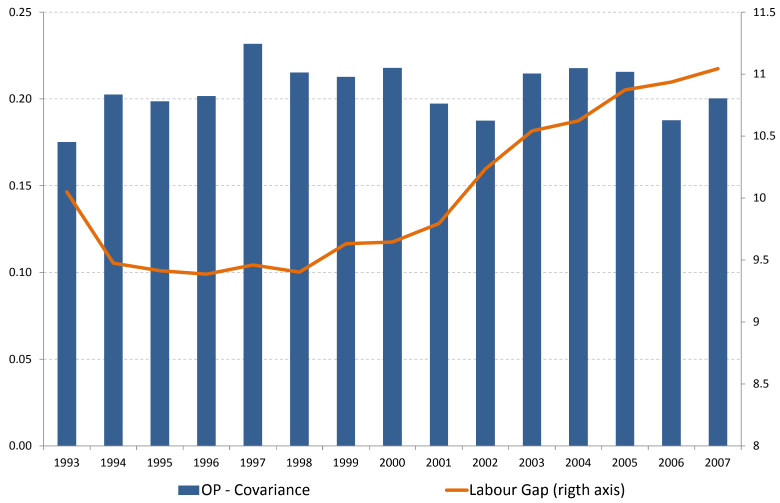
[43] We exclude observations with a growth rate of TFP variables – value added, fixed capital, material inputs above/below the 99th/1st percentile of the relative distribution. We also make sure that firm balance sheets covers 12 months. Results are robust to alternative thresholds.

[44] We limit the analysis to the manufacturing sector only to ease the interpretation of TFP estimation coefficients as marginal products; the underlying methodology however can be applied to other industries as well.

6.4. Measure Misallocation: Related Methods

An alternative technique to measure the degree of resource misallocation by industry is the already mentioned one pioneered by Olley and Pakes (1996) - OP. This methodology relies on a decomposition of sector-specific productivity in two main components: the "average" firm productivity and the covariance between firm size and productivity. The second terms aims to capture the efficiency of resource allocation within a given sector: covariance would be zero if production factors were randomly distributed[45].

Figure 14 – Industry Misallocation index (Olley and Pakes covariance term)



For sake of comparison with our approach, Figure 14 reports both the covariance based misallocation index (bars) as well as the marginal return to cost gaps (line). The OP term is computed considering the covariance between labor productivity (real value added per employees) and firm size. Industry values are then aggregated for the overall manufacturing sector using the sectoral value added share as weight[46].

[45]In detail, industry productivity Ω_t can be written as $\Omega_t = \sum_i \theta_{it}\omega_{it} = \overline{\omega_{it}} + \sum_i \left(\theta_{it} - \overline{\theta_t}\right)\left(\omega_{it} - \overline{\omega_t}\right)$, where ω_{it} is the productivity of firm i and θ_{it} a measure of its relative size within the sector (employment share). Aggregate industry productivity, Ω_t, is then equal to the average firm productivity $\overline{\omega_{it}}$ plus the covariance between productivity and economic size, aiming to capture the strength of the relation between productivity and the market share.

[46]For comparison purposes we follow the same procedure as Berthou and Sandoz (2014): defining productivity as value added per employees and θ as firm (industry) value added share; note that the covariance term tend to

Interestingly, the two indicators seem to move in the same direction, pointing at a relative increase in resource allocation efficiency up until 1997 (covariance raise and decline of the average labor gap). From this point on, optimal allocation deteriorate at the industry level (covariance) as well as the firm level (labor gap). Notably, both indicators qualify the period 2000-2002 as particularly detrimental. For the last years industry covariance seems to improve while at the micro-level the overall trend appears only to slow down.

be higher if we follow Bartelsman et al. (2013). Where the latter defines productivity as revenues (deflated) per employees and the weighting scheme is based on employment shares. Following this alternative specification we find almost the same values as in Bartelsman et al. (2013), around 30%.